**Covid-19 Big Data Analysis**

The aim of this project was to analyze the global spread of COVID-19 using a dataset containing daily confirmed case counts for various countries and regions. The analysis focused on visualizing trends, identifying the maximum infection rates for countries, and correlating these rates with socio-economic factors such as GDP per capita, social support, healthy life expectancy, and freedom to make life choices.

*The specific aims of this project includes :*

1. Analyze Global Trends in COVID-19 Spread
2. Calculate Maximum Infection Rates
3. Correlate Infection Rates with Socio-Economic Factors
4. Visualize Data Insights

**RESULTS :**

COVID-19 Trends:

* Visualized the progression of COVID-19 in various countries, such as China, Italy, and Spain, using time-series plots.
* Identified significant differences in infection trends across countries

Maximum Infection Rates:

* Calculated the maximum daily increase in confirmed cases for each country, highlighting the most critical points of infection spread (e.g., China: 15,136 cases in a single day, Italy: 6,557 cases).

Correlation Analysis:

* Joined the COVID-19 data with the World Happiness Report to analyze the relationship between maximum infection rates and socio-economic factors.
* Found positive correlations between maximum infection rates and factors like GDP per capita and healthy life expectancy, though correlations were relatively weak.

Visualization:

Created scatterplots and regression plots to visualize relationships:

* GDP per Capita vs. Maximum Infection Rates: Higher GDP per capita was moderately associated with higher maximum infection rates.
* Social Support vs. Maximum Infection Rates: Showed a weak positive correlation.
* Healthy Life Expectancy and Freedom to Make Life Choices vs. Maximum Infection Rates: Both factors displayed weak associations.